This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
G01N 33/68, 33/543, C07K 17/00
C12Q 1/70, 1/68

A1

(11) International Publication Number: WO 93/22684

(43) International Publication Date: 11 November 1993 (11.11.93)

(21) International Application Number: PCT/US93/04145

(22) International Filing Date: 28 April 1993 (28.04.93)

(30) Priority data:
07/876,792 29 April 1992 (29.04.92) US

(60) Parent Application or Grant
(63) Related by Continuation
US 07/876,792 (CIP)
Filed on 29 April 1992 (29.04.92)

(71) Applicant *(for all designated States except US)*: AFFYMAX TECHNOLOGIES N.V. [NL/NL]; De Ruyderkade 62, Cuaracao (AN).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : FODOR, Stephen, P., A. [US/US]; 3863 Nathan Way, Palo Alto, CA 94303 (US). STRYER, Lubert [US/US]; 843 Sonoma Terrace, Stanford, CA 94305 (US).

(74) Agents: WEAVER, Jeffrey, K. et al.; Townsend and Townsend Khourie and Crew, One Market Plaza, 20th Fl., Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: FACTORIAL CHEMICAL LIBRARIES

(57) Abstract

A method and library for determining the sequence of monomers in a polymer which is complementary to a receptor. The method provides for formation of pooled (6) and separate (10, 12) products. Separate products are subjected only to subsequent pooled coupling steps. Each pooled product is subsequently divided for formation of pooled and separate products. The resulting polymer library includes groups of polymer products. A first group of products (42) is used to identify the monomer at a first location in a polymer that is complementary to a receptor. A second group of products (44) is used to identify the monomer at a second location in a polymer that is complementary to a receptor.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

FACTORIAL CHEMICAL LIBRARIES

BACKGROUND OF THE INVENTION

The present invention relates to the field of polymer screening. More specifically, in one embodiment the invention provides an improved polymer library and method of using the library to identify a polymer sequence that is complementary to a receptor.

Many assays are available for measuring the binding affinity of receptors and ligands, but the information which can be gained from such experiments is often limited by the number and type of ligands which are available. Small peptides are an exemplary system for exploring the relationship between structure and function in biology. When the twenty naturally occurring amino acids are condensed into peptides they form a wide variety of three-dimensional configurations, each resulting from a particular amino acid sequence and solvent condition. The number of possible pentapeptides of the 20 naturally occurring amino acids, for example, is 20^5 or 3.2 million different peptides. The likelihood that molecules of this size might be useful in receptor-binding studies is supported by epitope analysis studies showing that some antibodies recognize sequences as short as a few amino acids with high specificity.

Prior methods of preparing large numbers of different oligomers have been painstakingly slow when used at a scale sufficient to permit effective rational or random screening. For example, the "Merrifield" method, described in Atherton et al., "Solid Phase Peptide Synthesis," IRL Press, (1989), incorporated herein by reference for all purposes, has been used to synthesize peptides on a solid support such as pins or rods. The peptides are then screened to determine if they are complementary to a receptor. Using the Merrifield method, it is not economically practical to screen more than a few peptides in a day.

Similar problems are encountered in the screening of other polymers having a diverse basis set of monomers. For example, various methods of oligonucleotide synthesis such as the phosphite-triester method and the phosphotrieseter method, described in Gait, "Oligonucleotide Synthesis," IRL Press, (1990), incorporated herein by reference for all purposes, have similar limitations when it is desired to synthesize many diverse oligonucleotides for screening.

To screen a larger number of polymer sequences, more advanced techniques have been disclosed. For example, Pirrung et al., WO 90/15070, incorporated herein by reference for all purposes, describes a method of synthesizing a large number of polymer sequences on a solid substrate using light directed methods. Dower et al., U.S. application Serial No. 07/762,522, also incorporated by reference herein for all purposes, describes a method of synthesizing a library of polymers and a method of use thereof. The polymers are synthesized on beads, f r example. A first monomer is attached to a pool of beads. Thereafter, the pool of beads is divided, and a

second mon mer is attached. The process is repeated until a desired, diverse set of polymers is synthesized.

Other methods of synthesizing and screening polymers have also been proposed. For example, Houghten et al., "Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery," Nature (1991) 354:84-86, discuss a method of generating peptide libraries that are used for screening the peptides for biological activity (see also, Houghton et al., "The Use of Synthetic Peptide Combinatorial Libraries for the Identification of Bioactive Peptides," Peptide Research (1992) 5:351-358). Houghten synthesized a peptide combinatorial library (SPCL) composed of some 34 x 10^6 hexapeptides and screened it to identify antigenic determinants that are recognized by a monoclonal antibody. Furka et al., "General Method for Rapid Synthesis of Multicomponent Peptide Mixtures," Int. J. Peptide Protein Res. (1991) 37:487-493, discusses a method of synthesizing multicomponent peptide mixtures. Furka proposed pooling as a general method for the rapid synthesis of multicomponent peptide mixtures and illustrated its application by synthesizing a mixture of 27 tetrapeptides and 180 pentapeptides. Lam et al., "A new type of synthetic peptide library for identifying ligand-binding activity," Nature (1991) 354:82-84 used pooling to generate a pentapeptide bead library that was screened for binding to a monoclonal antibody. Blake et al. "Evaluation of Peptide Libraries: An Interative Strategy To Analyze the Reactivity of Peptide Mixtures With Antibodies," Bioconjugate Chem. (1992) 3:510-513 discusses the screening of presumed mixtures of 50,625 tetrapeptides and 16,777,216 hexpeptides to select epitopes recognized by specific antibodies.

Lam's synthetic peptide library consists of a large number of beads, each bead containing peptide molecules of one kind. Beads that bind a target (e.g., an antibody or strepavidin) are rendered colored or fluorescent. Lam reports that several million beads distributed in 10-15 petri dishes can be screened with a low-power dissecting microscope in an afternoon. Positive beads are washed with 8M guanidine hydrochloride to remove the target protein and then sequenced. The 100-200 μm diameter beads contain 50-200 pmol of peptide, putatively well above their 5 pmol sensitivity limit. Three pentapeptide beads were sequenced daily. The essence of Lam's method is that the identity of positive beads is established by direct sequencing.

Houghten et al. use a different approach to identify peptide sequences that are recognized by an antibody. Using the nomenclature described herein, Houghten et al. screened an $X_6X_5X_4X_3X_2X_1$ library and found that the mixture $DVX_4X_3X_2X_1$ had greatest potency in their inhibition assay. Houghten then synthesized a $DVX_4X_3X_2X_1$ library and identified the most potent amino acid in the third position. After three more iterations, they found that DVPDYA binds to the antibody with a K_d of 30 nM. The essence of Houghten's method is recursive retrosynthesis, in which th number of pooled positions decreases by one each iteration.

Blak _t al. used a "bogus coin strategy" to guide them to a preferred amino acid sequence. In this strategy a basis set of mon mers (15 amino acids) is first divided into three

groups. Blake et al. ch se A, L, V, F, Y (subgroup α), G, S, P, D, E (subgroup β), and K, R, H, N, Q (subgroup γ). By adjusting the "weighting of the subgroups at each position in the polymer sequence, and then testing the activity of the weighted polymer against an unweighted polymer, one subgroup was selected for each monomer position in the sequence. In an experiment conducted by Blake et al., a complete collection of tetramers X_{1P} X_{2P} X_{3P} X_{4P} was reduced to α_1 α_2 γ_3 α_4 by four inhibition experiments. Then the subgroups α and γ were each further subdivided into three groups of amino acid which were used to synthesize four more collections of weighted polymers. Inhibition studies with each of these collections suggested an epitope (F or Y)$_1$ (A or L)$_2$ (K or R)$_3$ (F or Y)$_4$. One more iteration gave the desired epitope FLRF.

While meeting with some success, prior methods have also met with certain limitations. For example, it is sometimes desirable to avoid the use of the equipment necessary to conduct light directed techniques. Also, some prior methods have not produced the desired amount of diversity as efficiently as would be desired.

From the above, it is seen that an improved method and apparatus for synthesizing a diverse collection of chemical sequences is desired.

SUMMARY OF THE INVENTION

An improved polymer library and method of screening diverse polymers is disclosed. The system produces libraries of polymers in an efficient manner, and utilizes the libraries for identification of the monomer sequence of polymers which exhibit significant binding to a ligand.

According to one aspect of the invention, a library of polymers is formed using "pooled" and "unpooled" (or "separate") coupling steps. In the pooled steps, each of the monomers from a basis set of monomers is coupled to the terminus of a growing chain of monomers on a plurality of previously mixed solid substrates. The mixed substrates are divided for coupling of each individual monomer in a basis set. In separate steps, the substrates are not intermixed from a previous coupling step, and each of the monomers in a basis set is separately coupled to the terminus of a growing chain of monomers on a plurality of the unmixed substrates.

According to one preferred aspect of the invention, pooled steps and unpooled steps are ordered such that the identification of a monomer sequence which binds to a receptor can be readily identified from the library. For example, according to one preferred embodiment of the invention, several groups of products are derived from the synthesis steps. Each group is used to identify the monomer at a specific position in the polymer chain.

According to most preferred aspects of the invention, the library is constructed using an ordered series of coupling st ps in which products resulting from a separate step are, thereafter, only subjected to pooled coupling steps. Products resulting from a pooled coupling step which have n t been previously subjected to an unpooled step are always divided for pooled and

unpooled coupling. This ordered series of steps results in a relatively small number of coupling steps, but still allows for identification of the monomer sequence of a polymer which is complementary to a receptor of interest. For example, a first group of products is used to identify the monomer at a first location in a polymer that is complementary to a receptor. A second group of products is used to identify the monomer at a second location in a polymer that is complementary to a receptor.

Accordingly, in one embodiment of the invention provides a polymer library screening kit. The kit includes families of polymers X_3-X_{2p}-X_{1p}, X_{3p}-X_2-X_{1p}, and X_{3p}-X_{2p}-X_1 wherein X_{3p}-X_{2p}-X_1 comprises a collection of at least first and second polymer mixtures, the first polymer mixture having a first monomer in a first position of polymer molecules therein, and different monomers in second and third positions of the polymer molecules therein, and wherein the second polymer mixture has a second monomer in the first position of polymer molecules therein, and different monomers in second and third positions of the polymer molecules therein; X_{3p}-X_2-X_{1p} comprises a collection of at least third and fourth polymer mixtures, the third polymer mixture having a third monomer in the second position and the fourth polymer mixture having a fourth monomer in the second position, each of the third and fourth polymer mixtures having different monomers in the first and third positions; and X_3-X_{2p}-X_{1p} comprises a collection of at least fifth and sixth polymer mixtures, the fifth polymer mixture having a fifth monomer in the third position and the sixth polymer mixture having a sixth monomer in the third position, each of the fifth and sixth polymer mixtures having different monomers in the first and second positions, wherein the first, third, and fourth monomers are the same or different and the second, fourth, and fifth monomers are the same or different.

A method of identifying first and second monomers in a polymer that is complementary to a receptor is also provided. The method includes the steps of coupling first and second monomers in a first basis set to individual substrates and mixing substrates to form first pooled products; coupling the first and second monomers from the first basis set to individual substrates, and not mixing the substrates to form at least first and second separate products; separately coupling first and second monomers from a second basis set to substrates from the first pooled products and not mixing the substrates to form at least third and fourth separate products, the second basis set being the same or different than the first basis set; coupling the first and second monomers from the second basis set to individual substrates from the first separate products and mixing the substrates to form second pooled products; coupling the first and second monomers from the second basis set to individual substrates from the second separate products to form third pooled products; and exposing a receptor to the third and fourth separate products to identify a second monomer in a polymer which is complementary to a receptor, and exposing the second and third pooled products to the receptor to identify a first monomer in a polymer which is complementary to a receptor.

A polymer screening technique using factoring is also disclosed.

A further understanding of the nature and advantages of the inventions herein may be realized by reference to the remaining portions f the specificati n and the attached drawings.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1a and 1b are schematic diagrams of specific embodiments of the invention;

Fig. 2 illustrates a simple reacti n graph;

Fig. 3 illustrates a reaction graph with pooled and separate products;

Fig. 4 illustrates a simplified reaction graph;

Figs. 5a, 5b, and 5c illustrate a family of pooled syntheses;

Fig. 6 illustrates a reaction graph for forming the products $X_{3p}X_2X_{1p}$;

Fig. 7 illustrates a reaction graph for all 64 trinucleotides;

Fig. 8 illustrates the synthesis of AAT, TGC, TGT, GTA, GTG, and CCG;

Fig. 9 provides an alternative representation of the invention;

Figs. 10a, 10b, and 10c illustrate a recursive retrosynthesis embodiment of the invention;

Figs. 11a, 11b, and 11c illustrate a combinatorial synthesis chamber of the invention; and

Figs. 12 illustrates a polymer library according to one embodiment of the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

CONTENTS

I. Terminology

II. Overall Description

III. Polynomial Factoring Applied to Screening

IV. Conclusion

I. Terminology

Ligand: A ligand is a molecule that is recognized by a particular receptor. Examples of ligands that can be investigated by this invention include, but are not restricted to, agonists and antagonists for cell membrane receptors, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), hormone receptors, peptides, enzymes, enzyme substrates, cofactors, drugs, lectins, sugars, oligonucleotides, nucleic acids, oligosaccharides, proteins, and monoclonal antibodies.

M nomer: A member of the set of small m lecules which are r can be joined togeth r to form a polymer. The set of monomers includes but is not restricted to, for example, the set of comm n

L-amin acids, the set f D-amin acids, the set f synthetic and/or natural amino acids, th set of nucleotides and the set of pentoses and hexoses, as well as subsets thereof. The particular ordering of monomers within a polymer is referred to herein as the "sequence" of th polymer. As used herein, monomers refers to any member of a basis set for synthesis of a polymer. For example, dimers of the 20 naturally occurring L-amino acids form a basis set of 400 monomers for synthesis of polypeptides. Different basis sets of monomers may be used at successive steps in the synthesis of a polymer. Furthermore, each of the sets may include protected members which are modified after synthesis. The invention is described herein primarily with regard to the preparation of molecules containing sequences of monomers such as amino acids, but could readily be applied in the preparation of other polymers. Such polymers include, for example, both linear and cyclic polymers of nucleic acids, polysaccharides, phospholipids, and peptides having either α-, β-, or ω-amino acids, heteropolymers in which a known drug is covalently bound to any of the above, polynucleotides, polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, polyacetates, or other polymers which will be apparent upon review of this disclosure. Such polymers are "diverse" when polymers having different monomer sequences are formed at different predefined regions of a substrate. Methods of cyclization and polymer reversal of polymers which may be used in conjunction with the present invention are disclosed in copending application Serial No. 796,727, filed November 22, 1991 entitled "POLYMER REVERSAL ON SOLID SURFACES," incorporated herein by reference for all purposes. The "position" of a monomer in a polymer refers to the distance, by number of monomers, from a terminus or other reference location on a polymer.

Peptide: A polymer in which the monomers are alpha amino acids and which are joined together through amide bonds, alternatively referred to as a polypeptide. In the context of this specification it should be appreciated that the amino acids may be the L-optical isomer or the D-optical isomer. Peptides are often two or more amino acid monomers long, and often more than 20 amino acid monomers long. Standard abbreviations for amino acids are used (e.g., P for proline). These abbreviations are included in Stryer, Biochemistry, Third Ed., 1988, which is incorporated herein by reference for all purposes.

Receptor: A molecule that has an affinity for a given ligand. Receptors may be naturally-occurring or manmade molecules. Also, they can be employed in their unaltered state or as aggregates with other species. Receptors may be attached, covalently or noncovalently, to a binding member, either directly or via a specific binding substance. Examples of receptors which can be employed by this inventi n include, but are not restricted to, antibodies, cell membrane receptors, m noclonal antibodies and antisera reactive with specific antigenic determinants (such as

on viruses, cells r other materials), drugs, polynucleotides, nucleic acids, peptides, cofactors, lectins, sugars, polysaccharides, cells, cellular membranes, and organelles. Receptors are sometimes referred to in the art as anti-ligands. As the term receptors is used herein, no difference in meaning is intended. A "Ligand Receptor Pair" is formed when two macromolecules have combined through molecular recognition to form a complex.

Specific examples of receptors which can be investigated by this invention include but are not restricted to:

a) Microorganism receptors: Determination of ligands which bind to receptors, such as specific transport proteins or enzymes essential to survival of microorganisms, is useful in a new class of antibiotics. Of particular value would be antibiotics against opportunistic fungi, protozoa, and those bacteria resistant to the antibiotics in current use.

b) Enzymes: For instance, the binding site of enzymes such as the enzymes responsible for cleaving neurotransmitters; determination of ligands which bind to certain receptors to modulate the action of the enzymes which cleave the different neurotransmitters is useful in the development of drugs which can be used in the treatment of disorders of neurotransmission.

c) Antibodies: For instance, the invention may be useful in investigating the ligand-binding site on the antibody molecule which combines with the epitope of an antigen of interest; determining a sequence that mimics an antigenic epitope may lead to the development of vaccines of which the immunogen is based on one or more of such sequences or lead to the development of related diagnostic agents or compounds useful in therapeutic treatments such as for autoimmune diseases (e.g., by blocking the binding of the "self" antibodies).

d) Nucleic Acids: Sequences of nucleic acids may be synthesized to establish DNA or RNA binding sequences.

e) Catalytic Polypeptides: Polymers, preferably polypeptides, which are capable of promoting a chemical reaction involving the conversion of one or more reactants to one or more products. Such polypeptides generally include a binding site specific for at least one reactant or reaction intermediate and an active functionality proximate to the binding site, which functionality is capable of chemically modifying the bound reactant. Catalytic polypeptides and others are described in, for example, PCT Publication No. WO 90/05746, WO 90/05749, and WO 90/05785, which are incorporated herein by reference for all purposes.

f) Hormone receptors: For instance, the receptors for insulin and growth hormone. Determination of the ligands which bind with high affinity to a receptor is useful in the development of, for example, an oral replacement of the daily injections which diabetics must tak to relieve the symptoms of diabetes, and in the other case, a replacement for the scarce human growth horm ne which can only be obtained from cadavers or by recombin-

ant DNA technology. Other xamples are the vasoconstrictive hormon receptors; determination of those ligands which bind to a receptor may lead t the development of drugs t control blood pressure.

g) Opiate receptors: Determination of ligands which bind to the opiate receptors in the brain is useful in the development of less-addictive replacements for morphine and related drugs.

Substrate or Solid Support: A material having a surface and which is substantially insoluble in a solution used for coupling of monomers to a growing polymer chain. Such materials will preferably take the form of small beads, pellets, disks or other convenient forms, although other forms may be used. A roughly spherical or ovoid shape is preferred.

Basis Set: A group of monomers that is selected for attachment to a solid substrate directly or indirectly in a given coupling step. Different basis sets or the same basis sets may be used from one coupling step to another in a single synthesis.

Synthetic: Produced by in vitro chemical or enzymatic synthesis. The synthetic libraries of the present invention may be contrasted with those in viral or plasmid vectors, for instance, which may be propagated in bacterial, yeast, or other living hosts.

Symbols

x_i denotes the set of monomer units in reaction round i.

x_{ij} denotes the j'th monomer unit in reaction round i; x_{ij} can be a null (∅).

S_i refers to the separated products after reaction round i.

P_i refers to the pooled products of round i and all preceding rounds.

X_{ip} denotes the pooling of reactants of round i only

Reaction Graphs

A filled circle • denotes a reaction product terminating in a particular monomer unit x_{ij}. The set of reaction products terminating in x_i is shown by a set of circles on the same horizontal level.

Filled circles that react with each other are connected by straight lines. Pooling is shown by lines meeting below in an open circle.

A factorable polynomial synthesis is one in which each monomer unit of a round is joined to each monomer of the preceding round. In a graph of such a synthesis, each filled circl at on level is connected to each filled circl of the lev l above. For example, the reaction graph corresponding to a three-round factorable synthesis with

$$X_1 = X_2 = X_3 = \{A,T,G,C\}$$

which yields all 64 trinucleotides, is sh wn in Fig. 7.

In contrast, in an irreducible (prime) polynomial synthesis, at least one line in the graph of the corresponding factorable polynomial synthesis is missing. In the synthesis of AAT, TGC, TGT, GTA, GTG, and CCG only, such syntheses are illustrated in Fig. 8.

II. Overall Description

Fig. 1 is an overall illustration of one aspect of the invention. As shown therein, monomers A and B, which form all or part of a first basis set of monomers, are coupled to substrates 2 in vessels 4a and 4b. The substrates in each of the vessels 4a and 4b are divided. A portion of the substrates from each of vessels 4a and 4b are mixed in vessel 6, and divided for a subsequent coupling step into vessels 6a and 6b. Another fraction of the monomers from vessels 4a and 4b is not mixed, as indicated by vessels 10 and 12.

Thereafter, the substrates are coupled to monomers from a second basis set C,D, which may or may not be the same as the basis set A,B. As shown, the monomer C is coupled to the mixed or "pooled" substrates in vessel 6a, while the monomer D is coupled to the "pooled" substrates in vessel 6b. A portion of the products of these reactions may be mixed for later coupling steps, but at least a portion of the products in vessels 6a and 6b are not mixed.

The products in vessels 10 and 12 are preferably each divided for coupling to monomer C as shown in vessels 10b and 12b, while the substrates in vessels 10a and 12a are used to couple the monomer D to the growing polymer chain. The products of the reactions in vessels 10a and 10b are mixed or pooled, and placed in vessel 20. The products of the reactions in vessels 12a and 12b are mixed or pooled, and placed in vessel 22.

The products in vessels 20 and 22 are, thereafter, used to identify a first monomer in a polymer which is complementary to a receptor of interest. It is assumed for the sake of illustration herein that the monomer sequence AC is complementary to the receptor R. A receptor labeled with, for example, a fluorescent or radioactive label *, is exposed to the materials in vessels 20 and 22, and unbound receptor is separated from the solid supports. Binding to the substrates will occur only with the substrates in vessel 20. Fluorescence is, therefore, observed only in vessel 20. From this observation, it is possible to conclude that the first monomer in a complementary receptor is A, since all of the polymers in vessel 22 contain the first monomer B. Conversely, all of the polymers in vessel 20 contain the first monomer A.

The labeled receptor is also exposed to the polymers in vessels 26a and 26b. In this case, binding of th labelled receptor will be observed only in vessel 26a. Accordingly, it is possibl to identify the second monomer in a complementary sequence as C, since none of th polymers in vessel 26b contain th second mon mer C, whil all of the polymers in vessel 26a

contain the second mon mer C. Therefore, it is possible to conclude that the sequence AC is complementary to R since binding is observed in vessels 26a and 20.

Fig. 1b illustrates aspects of a preferred embodiment of the invention in greater detail with a larger polymer chain. According to the embodiment shown in Fig. 1b, a basis set of 3 monomers, A, B, and C is used in each coupling step. The synthesized polymers are to be three monomers long. It will be recognized by those of skill in the art that the number of monomers in a basis set and the number of coupling steps will vary widely from one application to another. Also, intervening coupling steps of, for example, common monomer sequences may be used in some embodiments. Therefore, when a polymer is represented by, for example, the notation "ABC" or "ABE" herein, it is to be understood that other common monomers may be added such that ABDC and ABDE are represented by ABC and ABE. The embodiment shown in Fig. 1b is provided merely as an illustration of the invention.

As shown in Fig. 1b, the synthesis takes place on a plurality of substrates 2. According to a preferred aspect of the invention, the substrates 2 take the form of beads, such as those made of glass, resins, plastics, or the like. The term "beads" is used interchangeably herein with the word "substrate," although it is to be understood that the beads need not take on a circular or ovoid shape and can take the form of any suitable substrate. It will be further understood that the substrates 2 are shown only in the top portion of Fig. 1b, but the substrates will be present in each of the reaction products shown in Fig. 1b to the left of the monomer sequences. In each vessel in Fig. 1b, all of the possible polymer products are listed. Many "copies" of each sequence will generally be present.

According to one embodiment, conventional Merrifield techniques are used for the synthesis of peptides, such as described in Atherton et al., "Solid Phase Peptide Synthesis," IRL Press, (1989), previously incorporated herein by reference for all purposes. Of course other synthesis techniques will be suitable when different monomers are used. For example, the techniques described in Gait et al., Oligonucleotide Synthesis, previously incorporated by reference herein by reference for all purposes, will be used when the monomers to be added to the growing polymer chain are nucleotides. These techniques are only exemplary, and other more advanced techniques will be used in some embodiments such as those for reversed and cyclic polymer synthesis disclosed in U.S. application Serial No. 07/796,727, previously incorporated herein by reference for all purposes.

A large number of beads are utilized such that the beads may be separated into separate reaction vessels in later steps and still be present in sufficient numbers such that the presence of a complementary receptor may be detected. As a general rule, it will be desired to use 10 to 100 or more times th number of combinatorial possibilities for the synthesis so as to ensure each member of each set is synthesized. Also, the use of a large number of beads ensures that

pooled reaction products are distributed to each succeeding reaction vessel when a pooled group of beads is divided.

The beads are preferably as small as possible so that the reaction vessels and other material handling equipment utilized in the process may also be as small as possible. Preferably, the beads have a diameter of less than about 1 mm, and preferably less than about 100 μm, and more preferably less than about 10 μm. In some embodiments, the synthesis is carried out in solution. In other embodiments, the synthesis is carried out on solid substrates, and the resulting polymers are then cleaved from the substrates before binding with a receptor.

As shown in Fig. 1b the monomers A, B, and C are coupled to substrates in three reaction vessels 4a, 4b, and 4c, respectively. A single substrate is shown in Fig. 1b for purposes of clarity, but it will be recognized that in each reaction vessel a large number of beads will be present. Accordingly, a large number of "copies" of the substrates with the respective monomers coupled thereto are formed in each of reaction vessels 4a, 4b, and 4c. It will be recognized that the monomers need not be directly coupled to the substrate, and in most cases linker molecules will be provided between the monomers and the substrate, such as those described in U.S. application Serial No. 07/624,120, incorporated herein by reference for all purposes. Also, it should be recognized that the steps shown in Fig. 1b may be preceded by or followed by other synthesis steps which may or may not be combinatorial steps using the techniques described herein.

Thereafter, a fraction of the products in each of vessels 4a, 4b, and 4c are combined, mixed, and redistributed to each of reaction vessels 6a, 6b, and 6c. The remaining fraction of the products in each of vessels 4a, 4b, and 4c is not combined. Instead, the remaining fraction of the products in reaction vessel 4a is divided and placed in reaction vessels 8a, 8b, and 8c. Similarly, the remaining fraction of the products in vessel 4b is divided and placed in vessels 10a, 10b, and 10c. The remaining fraction of the products in reactant vessel 4c is divided and placed in reaction vessels 12a, 12b, and 12c.

The reactants placed in vessels 6a, 6b, and 6c are referred to herein as "pooled" reactants since they comprise a mixture of the products resulting from the previous coupling step. The reactants placed in vessels 8, 10, and 12 by contrast are separate reactants since they are not mixtures of the products from the previous coupling steps. According to a preferred embodiment of the invention, after the reactants in vessels 8, 10, and 12 are subjected to a separate coupling step, they are subjected only to pooled coupling steps thereafter. Conversely, in each subsequent coupling step, the pooled reactants are subjected to a coupling step, and divided for subsequent separate and pooled coupling steps.

Preferably, the reactants are divided such that a greater fraction of the beads is distributed for pooled synthesis. For example, in Fig. 9, 4/5 of the beads would g to the first pooled group 905 whil 1/5 would go to the unpooled group 903.

Thereafter the monomers A, B, and C are coupled to the growing polymer chain in reaction vessels 8a, 8b, and 8c, respectively. The resulting polymers then have th monomer sequence CA, CB, and CC in reacti n vessels 8a, 8b, and 8c, respectively. The products of these reactions are then mixed or pooled in reaction vessel 9, and the mixture is again divided among reaction vessels 14a, 14b, and 14c. The monomers A, B, and C are again coupled to the growing polymer chains in vessels 14a, 14b, and 14c, respectively. The products of these reactions are again mixed or pooled and placed in vessel 16a.

Similarly, the monomers A, B, and C are coupled to the growing polymer chain in reaction vessels 10a, 10b, and 10c, then mixed in vessel 18, divided, and placed in reaction vessels 20a, 20b, and 20c. Monomers A, B, and C are coupled to the growing polymer chain in vessels 20a, 20b, and 20c respectively, mixed, and placed in vessel 16b. Monomers A, B, and C are also coupled to the growing polymer chain in reactant vessels 12a, 12b, and 12c respectively, mixed, and placed in vessel 21. These products are divided for reaction with monomers A, B, and C in vessels 22a, 22b, and 22c respectively, mixed, and placed in vessel 16c. A characteristic feature of the preferred embodiments of the present invention should be noted in the right half of Fig. 1b. Specifically, once the products of a reaction are not pooled (such as in vessels 8, 10, and 12), the products of coupling steps are always pooled thereafter.

Referring to the left hand portion of Fig. 1b, the pooled reactants in vessels 6a, 6b, and 6c are coupled to monomers A, B, and C respectively, resulting in the products shown in vessels 26a, 26b, and 26c. Since the products in vessels 26a, 26b, and 26c are derived from a "chain" of pooled reactions, the products are separated for both pooled and separate reactions. Specifically, a portion of the substrates in vessels 26a, 26b, and 26c are combined, mixed, and divided for pooled reactions with monomers A, B, and C in vessels 28a, 28b, and 28c respectively. In addition, the remaining portion of the products in vessels 26a, b, and c are separately divided and placed in reaction vessels 30a-c, 32a-c, and 34a-c respectively. The materials in vessels 30a, 32a, and 34a are coupled to monomer A, the materials in vessels 30b, 32b, and 34b are coupled to monomer B, and the materials in vessels 30c, 32c, and 34c are coupled to monomer C. Since the products in vessels 30, 32, and 34 result have been preceded by a separate reaction, the products in vessels 30, 32, and 34 are pooled, or mixed, and placed in vessels 36a, 36b, and 36c, respectively.

For reasons that will be discussed further below, the vessels in group 42 are used to determine the identity of the monomer in the first position in a polymer that is complementary to a receptor. The vessels in group 44 are used to determine the identity of the second monomer in a polymer that is complementary to a receptor. The vessels in group 46 are used to determine the identity f the third m nomer in a polymer that is complementary to a receptor.

For example, assume that a given recept r is complementary to the monomer sequence ABC, but the sequence of the complementary polymer is not known ab initio. If the

receptor is labelled with an appropriate label such as fluorescein and placed in each of th vessels in groups 42, 44, and 46, fluorescence will be detected only in vessels 16c, 36b, and 28c since the polymer sequence ABC appears only in these vessels. Fluorescence may be detected using, for example, the methods described in Mathies et al., U.S. Patent No. 4,979,824, incorporated herein in its entirety by reference for all purposes.

Since all of the polymers in vessel 16c have monomer A in the first position, and none of the polymers in vessels 16a or 16b have monomer A in the first position, it is readily determined that the monomer in the first position of a complementary polymer is the monomer A. Similarly, since all of the polymers in vessel 36b have the monomer B in the second position, it is readily determined that the monomer B must occupy the second position of a complementary polymer sequence. Similarly, since all of the polymers in vessel 28c have a C monomer in the third position, the complementary receptor must have a C in its third position. Therefore, it would readily be determined that the complementary sequence to the receptor has the monomer sequence ABC.

As will be seen upon careful examination of the sequences in the vessel groups 42, 44, 46, ambiguities will generally not arise, regardless of the monomer sequence which is complementary to the receptor of interest. As a point of comparison, if the receptor of interest is complementary to the sequence BBA, fluorescence would be detected only in vessels 16b, 36b, and 28a. From this information is becomes clear that the complementary monomer sequence must be BBA.

The above embodiment illustrates the synthesis of pooled groups of polymers by way of separation into separate vessels, followed by coupling and mixing. It will be recognized that this is only for convenience of illustration and that in some embodiments the pooled groups of polymers will be synthesized under controlled conditions by simultaneous reaction of each of the monomers to be coupled to the polymers in a single reactor. Further, it will be recognized that the synthesis steps above will be supplemented in many embodiments by prior, intermediate, and subsequent coupling steps, which are not illustrated for ease of illustration.

The above method may be generally illustrated by way of the adoption of appropriate nomenclature. For example, let X_i denote the set of monomer units that become joined to a growing chain at reaction round i. For example, suppose that

$$X_1=\{L,G\}\ X_2=\{P,Y\}\ X_3=\{R,A\}$$

A particular monomer is denoted by x_{ij}. For example,

$$x_{3,1} = R$$

The reaction products S_3 of such a three-round peptide synthesis is concisely represented by

$$S_3 = X_3X_2X_1$$

S_3 is determined by expanding a reaction polynomial as described in Fodor et al., Science (1991) 251:767-773, incorporated herein by reference for all purposes.

$$S_3 = (R+A)(P+Y)(L+G)$$

and so S_3 consists of 8 tripeptides:

RPL, RYL, RPG, RYG, APL, AYL, APG, and AYG

S_{ij} denotes a set of reaction products terminating in monomer unit x_{ij}. In the above synthesis, for example,

$$S_{12}=G \quad S_{21}=\{PL,PG\} \quad S_{32}=\{APL,AYL,APG,AYG\}$$

This three-round synthesis can also be represented by a reaction graph, as shown in Fig. 2. Each reaction product of round i is depicted by a filled dot on the same horizontal level. Each dot of round i is joined to each dot of the preceding round and to each dot of the succeeding round. For example, the dot denoting S_{21} is joined to the dots for S_{11} and S_{12}, and also to the dots S_{31} and S_{32}. Note that dots on a level are never connected to each other because, by definition, monomer units of a round do not combine with one another.

It is generally assumed that the products of each round are spatially separate and addressable. Each can then be readily assayed. However, the number of compounds generated by a combinatorial synthesis can, after a few rounds, greatly exceed the number of experimentally available bins or vessels. It is then advantageous to pool the products of one or more rounds of synthesis. For example, a five-round synthesis using the basic set of 20 amino acids yields 20^5 or 3.2×10^6 pentapeptides. In contrast, if the products of the first two rounds are pooled, the subsequent three rounds yield only 8,000 sets of products. Information is lost in the pooling process, but the number of products becomes experimentally tractable.

The above representation of combinatorial synthesis may be modified to take into account the effect of pooling. Suppose that products of the first two rounds of the three-round synthesis mentioned earlier are pooled. Th reaction graph for a with pooled steps is shown in Fig. 3. The pooled products of round i are denoted by P_i to distinguish them for the separate products

S_i. In a reaction graph, pooling is shown by the convergence of lines from the S_i that are pooled. P_i is then shown as an open circle.

In this example,

$$P_1 = \{L+G\} \quad P_2 = \{PL+PG+YL+YG\}$$

$$S_3 = X_3 P_2 = \{RPL+RPG+RYL+RYG,APL+APG+AYL+AYG\}$$

The plus sign joins products that are present in a mixture. In contrast, products separated by commas are located in separate bins and are spatially addressable. In this example, the pooled products of the second round are located in one bin, whereas the products after three rounds are located in two bins. One bin contains the mixture RPL+RPG+RYL+RYG, and the other bin contains the mixture APL+APG+AYL+AYG.

This reaction graph can be simplified. Suppose that P_1 was coupled to a equimolar mixture of x_{21} and x_{22} in a single bin. If the coupling efficiencies for all species are the same, the amounts and kinds of products obtained would be the same as that given by coupling P_1 with x_{21} and x_{22} in separate bins and then pooling the products. Thus, pooled products and pooled reactants are formally equivalent provided that the reactions occur in a substantially homogeneous solution and all coupling efficiencies are substantially the same. Hence, an X_3P_2 synthesis can be most simply represented by the reaction graph shown in Fig. 4.

The line joining P_2 to P_1 means that all products in P_1 are coupled equally to all reactants X_2, either by (1) adjusting the concentrations of reactants or (2) driving the reactions to completion in separate bins, followed by pooling. For beads or other discrete particles, (2) more often applies so that each particle expresses only one kind of product.

By way of comparison, the synthesis of 180 pentapeptides in Furka et al., "General Method for Rapid Synthesis of Multicomponent Peptide Mixtures," Int. J. Peptide Protein Res. (1991) 37:487-493, is represented with the above nomenclature as $S_5 = X_5P_4$, where $X_1=\{A\}$, $X_2=\{E,F,K,\}$, $X_3=\{E,P,K\}$, $X_4=\{E,F,G,K\}$, and $X_5=\{E,G,K,L,P\}$. The peptide combinatorial library synthesis in Houghten et al., "Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery," Nature (1991) 354:84-86 is $S_6 = X_6 X_5P_4$, where each X_i is a set of 18 naturally occurring amino acids. The S_6 products are located in 18x18 or 324 bins, each containing a mixture of $18^3 = 5,832$ hexapeptides. The pooled synthesis in Lam et al., "A new type of synthetic peptide library for identifying ligand-binding activity," Nature (1991) 354:82-84, is represented using the above nomenclature as P_5, where each X_i is a set of 19 naturally-occurring amino acids. P_5 is a mixture of $19^5 = 2.48\text{x}10^6$ beads, each bearing one kind of peptide.

In the pooled syntheses f Houghten, Lam, and Furka, all products from round 1 to round n are mixed. In Furka's synthesis (X_5P_4), the first f ur rounds are pooled. In an X_3P_2 synthesis, the first two rounds are pooled.

Representative pooled syntheses techniques according to one preferred embodiment of the invention herein are shown in Figs. 5a, 5b, and 5c. The symbol X_{ip} means that the reactants of round i have been pooled without pooling the reaction products of previous rounds. This is achieved by, for example, (1) mixing the reactants X_i or (2) by reacting each member of X_i with each reaction product of S_{i-1}, as shown in Fig. 6 for $X_{3p}X_2X_{1p}$.

For pentapeptides made of the naturally occurring 20 amino acids for example, a family of five pooled syntheses groups according to the invention herein will be particularly useful:

$$X_5X_{4p}X_{3p}X_{2p}X_{1p} \quad X_{5p}X_4X_{3p}X_{2p}X_{1p} \quad X_{5p}X_{4p}X_3X_{2p}X_{1p}$$

$$X_{5p}X_{4p}X_{3p}X_2X_{1p} \quad X_{5p}X_{4p}X_{3p}X_{2p}X_1$$

The products of each of these five syntheses product groups would be located in 20 physically isolated bins. Each bin would contain a different mixture of 160,000 pentapeptides. As with the trimer illustrated in Fig. 1b, the identity of the monomers forming a complementary pentamer would be determined unambiguously by identifying which of the 20 bins in each of the five syntheses product groups showed binding to a receptor.

It is to be recognized that while "bins" are referred to herein for the sake of simplicity, any of a variety of techniques may be used for physically separating the peptide or other polymer mixtures.

More specifically, a sequence of monomers in a complementary ligand for a receptor is identified as follows. For example, consider the family of pooled tripeptide libraries made of the 20 naturally occurring amino acids:

$$X_3X_{2p}X_{1p} \quad X_{3p}X_2X_{1p} \quad X_{3p}X_{2p}X_1$$

The most potent amino acid at the left position (x_{3i}) is revealed by analysis of the 20 bins of $X_3X_{2p}X_{1p}$; x_{2j} is determined by analysis of $X_{3p}X_2X_{1p}$; and x_{1k} is determined by analysis of $X_{3p}X_{2p}X_1$. The sequence of the most potent tripeptide is then predicted to be $x_{3i}x_{2j}x_{1k}$. Accordingly, each pooled group in the library reveals the identity of a monomer in a different position in a complementary polymer.

It will be recognized that it will not always be desirabl to determin the identity of the entire sequence of monomers in a polymer that is complementary to a receptor. Instead, it will only be necessary to d termine th identity of selected monomers in a polymer in some

instances. Th monomers of interest may be at intermediate locations on the chain of polymer, and may be interspersed by other monomers. Accordingly, in a more general sense, th method herein provides for the synthesis of a library of polymers. The library is used to identify at least two monomers of interest in the polymer chain.

For example, the identity of the x_{2j} monomer is determined by analysis of a library of polymers T-X_2-I-X_{1p}-T; and the identity of the monomer x_{1k} is determined by analysis of a library of polymers T-X_{2p}-I-X_1-T, where T indicates terminal groups on the polymer chain, which may be null groups, and I designates intermediate groups in the polymer chain, which may also be null groups.

The method of making the library used pooled and separate synthesis steps. The polymers have at least two monomer locations at which it is desired to determine the identity of monomers which provide a polymer with a sequence complementary to a receptor. The library is synthesized such that the products of a pooled synthesis are separated and subjected to a separate synthesis and a second pooled synthesis. The products of the separate synthesis are subjected to a series of pooled syntheses, without any further separate synthesis in preferred embodiments. Conversely, the products of the second pooled synthesis are divided and subjected to both a separate syntheses and a third pooled synthesis.

The synthesis steps result in a library of polymers having at least first and second subsets. The first subset is used to determine the identity of a monomer or monomers at a first location in the polymer chain which is complementary to a receptor. The second subset of the library is used to determine the identity of a monomer or monomers at a second location in the polymer chain which is complementary to a receptor.

The method uses summated assays to identify optimal sequences. The distribution of activities in the mixture assayed remains unknown. Only the aggregate activity is determined. More information can be obtained from analyses of beads or other particles that contain multiple copies of one kind of sequence. The activity of each bead can be quantitated even though its identity is unknown.

Suppose that 2 μm diameter beads are used for pooled syntheses. Some pertinent properties of typical beads are:

Volume = 4.2 μm^3

Surface area = 12.6 μm^2

Number of target sites = 1.3×10^5

(assuming 1 per 100 mm^2)

Number of beads per cm^3 = 2.4×10^{11}

Fluorescence measurements of beads flowing rapidly through a laser beam are made using techniques such as those in U.S. Patent N . 4,979,824, previously incorporated herein by reference for all purposes, which provide exemplary methods for determining the distribution of activities in a pooled synthesis.

Assume a light beam diameter of 2 μm is used for detection of fluorescein labeled beads, at a flow rate of 20 cm/s. The transit time of a bead through the beam is then 10 μs. The emission rate from a single chromophore can be as high as $10^8 s^{-1}$. If 10% of the target sites are occupied, this corresponds to an emission rate of about 10^{12} s^{-1}, or 10^7 emitted photons in 10 μs, which would be easily detected. If 10% of the sample volume is occupied by beads, an average of one bead would pass through the beam every 0.1 ms. Thus, 10^4 beads could be analyzed per second. A library of $3.2x10^6$ beads (each bearing a different pentapeptide) could be analyzed in about 6 minutes.

Alternatively, the beads may be analyzed by spreading them on a surface. For example, $3.2x10^6$ beads would occupy $1.28x10^7$ μm^2 if packed together in a square array. In 1.28 cm^2, these beads would occupy 10% of the surface area. Smaller beads, say 0.2 μm^2, would give a sufficient fluorescence signal. The advantage of smaller beads is that higher bead densities could be used, leading to a marked reduction in the time needed for analysis.

The fluorescence pulse height distribution emerging from either analysis would reveal whether there are many or few optimal sequences contained within the sample of beads. In the simplest case, a single bright bead is seen in just one bin of a pooled synthesis. The identity of the best sequence then comes directly from analysis of each pooled synthesis of the family.

In other cases, there is a distribution of intensities within several sets of beads. As a general rule, positioned libraries where binding is exhibited in multiple bins indicates that a particular position plays a less significant role in binding. In some embodiments, positions where ambiguity are detected are further evaluated through use of the VLSIPS™ technique. The VLSIPS™ arrays will vary only those positions wherein the monomer has not been determined unambiguously. The present invention is used, therefore, to reduce the number of polymers which will be screened with VLSIPS™ in some embodiments.

In still other cases, polymer mixtures synthesized in multiple bins are cleaved from their respective beads and then assayed for activity. The freed polymers are then able to interact with receptors in various orientations. The activity of such polymers can be assayed by various well-known techniques such as ELISA.

Fig. 9 provides an alternative description of the invention. As shown therein, at step 901 a collection of substrates is subjected to pooled and separate coupling steps, resulting in pooled and separate products 905 and 903, respectively. In comparison with th embodiment sh wn in Fig. 1b, products 903 are analogous to the products shown in vessels 8, 10, and 12, and products 905 are analogous to the products in vessel 6. The collection of substrate products 903

are then subjected to pooled coupling steps 903, 905, 907, and 909, i. ., the subsequent coupling steps to the separate reactants are nly pooled coupling steps. Accordingly, the identity of th monomer in the first position f a polymer complementary to a receptor is determined by evaluation of the products 907.

Conversely, the pooled products 905 are divided and subjected to pooled and separate coupling steps 909, resulting in pooled and separate products 907 and 913, respectively. As with the separate products 903, the separate products 913 are subjected only to pooled coupling steps thereafter, resulting in pooled products 915 and 917. The products 917 are used to determine the monomer in a second position in a polymer complementary to a receptor of interest.

In the same manner, the pooled products 907 are divided and subjected to pooled and separate coupling steps 919, resulting in pooled and separate products 923 and 921. The separate products 923 are subjected only to a pooled reaction thereafter, the products 925 being used to determine the monomer in a third position in a polymer complementary to a receptor of interest. The pooled products 921 are divided and subjected to pooled and separate reactions 927, resulting in pooled and separate products 929 and 931. The products 907, 917, 925, 931, and 929 are used to identify complementary receptors. In the preferred embodiment, the pooled products 927 are first used to determine if any polymers of interest are present. The separate products 931 are used to determine the identity of a monomer in a fourth position of a polymer complementary to a receptor.

As shown in Fig. 9, pooled products that have not been subjected to prior separate reactions are divided and subjected to pooled and separate reactions according to the invention herein. Conversely, products which result from a prior separate coupling step are only subjected to pooled coupling steps.

In an alternative embodiment depicted in Figs. 10a-10c, a recursive retrosynthesis is employed to screen a diverse set of polymers. Unlike the recursive retrosynthesis method of Houghten et al. described supra, this method identifies the sequence of a "best" polymer by identifying a collection of polymers ("library") containing the bead giving the strongest signal. Houghton et al., in contrast, identify the entire library, rather than the single polymer (bead), having the strongest signal. Thus, the technique of Houghton et al. may identify an incorrect monomer in the sequence of interest because of the library containing that monomer gave the strongest signal, while the "best" polymer is located in a different library. The recursive retrosynthesis embodiment of this invention overcomes this difficulty of Houghton et al.'s by identifying the individual polymer giving the strongest signal.

Referring to Figs. 10a-10c, an example of this process is described for the set of all pentamers formed from a basis set f 50 monomers. As shown in Fig. 10a, the complete collection of quadramers is synthesized on a number of beads (e.g., 3×10^6 beads) by four cycles f alternately dividing, reacting, and pooling th beads. Th pool of quadramers is then divided

into 50 bins, each of which is reacted with a different member of the basis set to give 50 bins of pentamers, each containing only those pentamers terminating in a specified monomer. In the notation used herein, this collection of polymers is represented by X_{1P} X_{2P} X_{4P} $X_{5(1-50)}$. The individual beads in each bin are then assayed to identify the single best bead (i.e., the bead providing the strongest signal on binding with the receptor of interest). This may be accomplished in about eight hours by FACS as described above, for example. Having determined the bin containing the bead providing the strongest signal, the identity of the monomer in the fifth position is known. In the example of Fig. 10, that monomer is D.

Next, the complete collection of trimers is formed as before (from e.g., 6×10^6 beads) by cycles of dividing, reacting, and pooling the beads as shown in Fig. 10b. Alternatively, the library of trimers could be set aside after the third cycle of the previous step (during formation of the complete library of pentamers). At this point, the pooled beads are divided into 50 bins, each of which is reacted with a different member of the basis set. The quadramers in each bin are then reacted with D to produce a collection of beads represented by X_{1P} X_{2P} X_{3P} $X_{4(1-50)}$ D. The beads in each bin are then assayed to again identify the bead giving the strongest signal. The bin from which that bead was taken identified the monomer at the next position: A in this case. The above process is repeated to produce X_{1P} X_{2P} $X_{3(1-50)}$ AD and identify the monomer at the next position as shown in Fig. 10c. In this example, the next monomer is identified as Q. The final two monomers of the sequence can be identified in the same fashion. However, it may be more efficient to simply screen the remaining 2,500 possible pentamers via a VLSIPS™ technique.

In general, for a polymer of length N synthesized from a basis set of n monomers, the terminal monomer may be identified by the following procedure. First, a pooled library of substrates is formed such that each substrate has a different polymer synthesized. The pooled library includes a collection of polymers represented by $X_{1p}X_{2p} \ldots X_{(N-1)p}$. The library is divided into n separate bins, each of which is then reacted with a different monomer to form a library $X_{1p}X_{2p} \ldots X_{(N-1)p}X_{N(1-n)}$. Finally, a receptor is exposed to the substrate in each of the n separate bins to identify the bin containing the polymer which binds the receptor most strongly. This bin provides the identify the monomer in the X_N position. The penultimate monomer is identified by a similar procedure from $X_{1p}X_{2p} \ldots X_{(N-1)(1-n)}R$, where R is the terminal monomer previously identified. Each succeeding monomer can be identified in the same manner. In this example, the two basis sets used to identify the monomers at positions N and N-1 each contained n members. It will be appreciated that the monomer basis sets used to identify the monomer at each position on the polymer may independent of the other basis sets.

A combinatorial synthesis chamber for conducting the synthesis, pooling, and dividing steps employed in each cycle of this invention is illustrated in Fig. 11a-11c. Individual chambers 200, each containing an amount of packed beads 203, are aligned in close proximity to one another to form a two-dimensional array. The reaction chambers 200 are mounted on a base

207 via passages 211. A filter 213 is provided at the base of each coupling chamber 200 to prevent the beads from leaving the coupling chambers 200 when solutions are drained through passages 211. In synthesis mode, coupling solutions are introduced through passages 205 while cover 201 is in a closed position (as shown in Fig. 11a). The contents of each chamber 200 are prevented from contacting adjacent chambers by cover 201. Passages 205 and 211 are controlled by valves or other control mechanisms not shown.

After the coupling reactions have proceeded to the desired extent, the coupling solutions are drained from the synthesis chamber through the passages 211. Subsequent washing steps may be necessary before pooling and redistribution. In such washing steps, washing solutions are introduced through passages 205 while cover 201 is held in the closed position. After sufficient time has elapsed, the washing solution is drained through passages 211. Multiple washing steps may be performed as necessary to remove unused coupling solutions from chambers 200. During the reacting and washing steps, the beads in the reaction chambers may be agitated by rotating or shaking the entire synthesis chamber.

As shown in Fig. 11b pooling and dividing of the beads is accomplished by sliding cover 201 away from base 207 to form a mixing chamber 215. The previously packed beads 203 are then agitated in a fluid such that they mix within mixing chamber 215. Ultimately, when the agitation is stopped, the beads will settle randomly into various chambers 200. At that point, the suspension solution can be drained through passages 211, and cover 201 can be lowered to rest on the tops of chamber 200 as shown in Fig. 11a. During the pooling and dividing stages, the valves in passages 205 and 211 are closed. A sealing means 209 is provided to prevent beads or fluid from leaving the synthesis chamber.

Fig. 11c shows a top view of a two-dimensional array of reaction chambers 200 mounted on base 207. The chambers shown are arranged in a 7 x 7 array of 49 chambers.

Fig. 12 illustrates a library of polymers which will be useful in accordance with the invention herein. As shown therein, the polymers have a number of monomer positions, designated by p_i. The monomers have at least two positions of interest, p_1 and p_2. p_1 and p_2 may in some embodiments be separated by various intermediate monomers or groups, and may also have various terminal groups attached thereto. The monomers are placed in a number of physically isolated bins or vessels 1002. The bins or vessels 1002 may in fact be attached, such as in a microtiter plate, or the bins/vessels may be distinct containers such as test tubes, microtiter trays, or the like.

A first bin 1002a contains polymers with a first monomer M_1 in the first position p_1 in each of the polymers therein. However, the polymer molecules in the first bin have a variety of different monomers such as M_1, M_2, and M_3 in a second position p_2. In the second bin 1002b a second monomer M_2 is in the first position p_1 in each of th polymers therein, while different m n mers such as M_1, M_2, and M_3 are in the second position p_2. In the third bin 1002c a third

monomer M_3 is in the first position p_1 in each f the polymers therein, while different m n mers such as M_1, M_2, and M_3 are in the second position p_2. The first, second, and third bins comprise all r part of a collection of bins ...X_1...X_2....

Conversely, fourth bin 1002d contains polymers with a first monomer M_1 in the second position p_2 in each of the polymer molecules therein. The polymer molecules in the first bin have a variety of different monomers such as M_1, M_2, and M_3 in their first position p_1. In the fifth bin 1002e a second monomer M_2 is in the second position p_2 in each of the polymers therein, while different monomers such as M_1, M_2, and M_3 are in the first position p_1. In the sixth bin 1002f a third monomer M_3 is in the second position p_2 in each of the polymers therein, while different monomers such as M_1, M_2, and M_3 are in the first position p_1. The fourth, fifth, and sixth bins comprise all or part of a collection of bins ...X_1...X_2....

In screening studies, the bins 1002a, 1002b, and 1002c are used to determine the identity of the monomer in position 1 of a polymer that is complementary to a receptor of interest. The bins 1002d, 1002e, and 1002f are used to determine the identity of the monomer in position 2 of a polymer that is complementary to a receptor of interest.

It will be recognized that the polymers which are screened according to the above methods can be of widely varying length and composition. For example, in preferred embodiments, the polymer molecules are preferably greater than 3 monomer units long, preferably greater than 5 monomer units long, more preferably greater than 10 monomer units long, and more preferably more than 20 monomer units long. Although a simplified library is shown in Fig. 12, it will be recognized that in most embodiments, the library will include additional polymer bins so as to identify the monomers at more than 3 positions, preferably more than 5 positions, more preferably more than 10 positions, and more preferably more than 20 positions in a complementary polymer to a receptor.

III. Polynomial Factoring Applied to Screening

In some embodiments a population of all possible polymers of length n are synthesized. If a receptor is found to bind with one of the polymers in the mixture, a second synthesis is conducted in which the polymers are "factored," i.e., two bins are formed, each having half of the population synthesized initially. It is then determined which of the two bins shows binding to the receptor, the bin which exhibits binding being referred to as a "target group." Yet another synthesis is conducted in which two bins are created, each with half of the population of the target group in the earlier bin. The process is repeated until the sequence of the polymer or polymers that show binding to the receptors is determined.

M re specifically, the inventi n provides for the synthesis of a population:

$$P = \sum^{n} X_i \quad \sum^{n} X_j$$

$$i=1 \quad j=1$$

This solution is factored as:

$$P = \sum_{i} X_i \left[\sum_{j}^{n/2} X_j + \sum_{n/2}^{n} X_j \right] = P_1 + P_2$$

where:

$$P_1 = \sum_{i=1}^{n} X_i \sum_{j=1}^{n/2} X_j \text{ ; and}$$

$$P_2 = \sum_{i=1}^{n} X_i \sum_{n/2}^{n} X_j$$

If P_1 generates a "hit," P_1 is factored. If P_2 generates a "hit," P_2 is factored. Each synthesis requires only half the number of polymers made in the prior step.

IV. Conclusion

The above description is illustrative and not restrictive. Many variations of the invention will become apparent to those of skill in the art upon review of this disclosure. Merely by way of example while the invention is illustrated primarily with regard to the synthesis of oligonucleotides and peptides, the invention will also find utility in conjunction with the synthesis and analysis of a wide variety of additional polymers. The scope of the invention should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the appended claims along with their full scope of equivalents.

WHAT IS CLAIMED IS:

1. A polymer library screening kit comprising families of polymers X_3-$X_{2p}X_{1p}$, X_{3p}-X_2-X_{1p} and X_{3p}-X_{2p}-X_1 wherein:

X_{3p}-X_{2p}-X_1 comprises a collection of at least first and second polymer mixtures, said first polymer mixture having a first monomer in a first position of polymer molecules therein, and different monomers in second and third positions of said polymer molecules therein, and wherein said second polymer mixture has a second monomer in said first position of polymer molecules therein, and different monomers in second and third positions of said polymer molecules therein;

X_{3p}-X_2-X_{1p} comprises a collection of at least third and fourth polymer mixtures, said third polymer mixture having a third monomer in said second position and said fourth polymer mixture having a fourth monomer in said second position, each of said third and fourth polymer mixtures having different monomers in said first and third positions; and

X_3-X_{2p}-X_{1p} comprises a collection of at least fifth and sixth polymer mixtures, said fifth polymer mixture having a fifth monomer in said third position and said sixth polymer mixture having a sixth monomer in said third position, each of said fifth and sixth polymer mixtures having different monomers in said first and second positions, wherein said first, third, and fourth monomers are the same or different and said second, fourth, and fifth monomers are the same or different.

2. The polymer library as recited in claim 1 wherein each of said polymer mixtures are selected from the group consisting of mixtures of peptides and mixtures of oligonucleotides.

3. The polymer library as recited in claim 1 wherein said polymers comprise at least four monomers.

4. The polymer library as recited in claim 1 further comprising labelled receptor molecules, and means for identifying mixtures of said polymers to which said receptor molecules are bound.

5. The polymer library as recited in claim 9 wherein said receptor molecules are labelled with a fluorescein label.

6. The polymer library as recited in claim 1 wherein said polymers are coupled to a solid substrate.

7. A method of identifying first and second monomers in a polymer that is complementary to a receptor comprising the steps of:

coupling first and second monomers in a first basis set to individual substrates and mixing substrates to form first pooled products;

coupling said first and second monomers from said first basis set to individual substrates, and not mixing said substrates to form at least first and second separate products;

separately coupling first and second monomers from a second basis set to substrates from said first pooled products and not mixing said substrates to form at least third and fourth separate products;

coupling said first and second monomers from said second basis set to individual substrates from said first separate products and mixing said substrates to form second pooled products;

coupling said first and second monomers from said second basis set to individual substrates from said second separate products to form third pooled products; and

exposing a receptor to said third and fourth separate products to identify a second monomer in a polymer which is complementary to a receptor, and exposing said second and third pooled products to said receptor to identify a first monomer in a polymer which is complementary to said receptor.

8. The method as recited in claim 7 wherein said step of exposing to a receptor is preceded by the step of performing additional steps of coupling and mixing to said second pooled products and said third pooled products.

9. The method as recited in claim 7 further comprising the step of mixing a portion of said third and fourth separate products to form fourth pooled products.

10. The method as recited in claim 9 further comprising the step of separately coupling monomers from a third basis set to said fourth pooled products.

11. The method as recited in claim 7 wherein said monomers are amino acids.

12. The method as recited in claim 7 wherein said monomers are nucleotides.

13. The method as recited in claim 7 wherein said steps are repeated to screen polymers having at least three monomers therein.

14. Th method as recited in claim 7 wherein at least one of said first and second mon mers cannot be determined unambigu usly, further comprising th steps of:

synthesizing an array of polymers, said potential complementary polymers using a light-directed synthesis technique; and

detecting binding of said receptor to said potential complementary polymers.

15. A library of polymers to be used for identification of a receptor complementary to at least one of said polymers comprising:

a first set of polymers, said first set of polymers having a first monomer in a first position, and a plurality of different monomers at a second position; and

a second set of polymers, isolated from said first set of polymers, said second set of polymers having a second monomer in said second position, and a plurality of different monomers in said first position.

16. A library of polymers as recited in claim 15 further comprising:

a third set of polymers, isolated from said first and second sets of polymers, said third set of polymers having a third monomer in said first position and a plurality of different monomers at said second position.

17. A library of polymers as recited in claim 16 further comprising:

a fourth set of polymers, isolated from said first, second, and third sets of polymers, said fourth set of polymers having a fourth monomer in said second position and a plurality of different monomers in said first position.

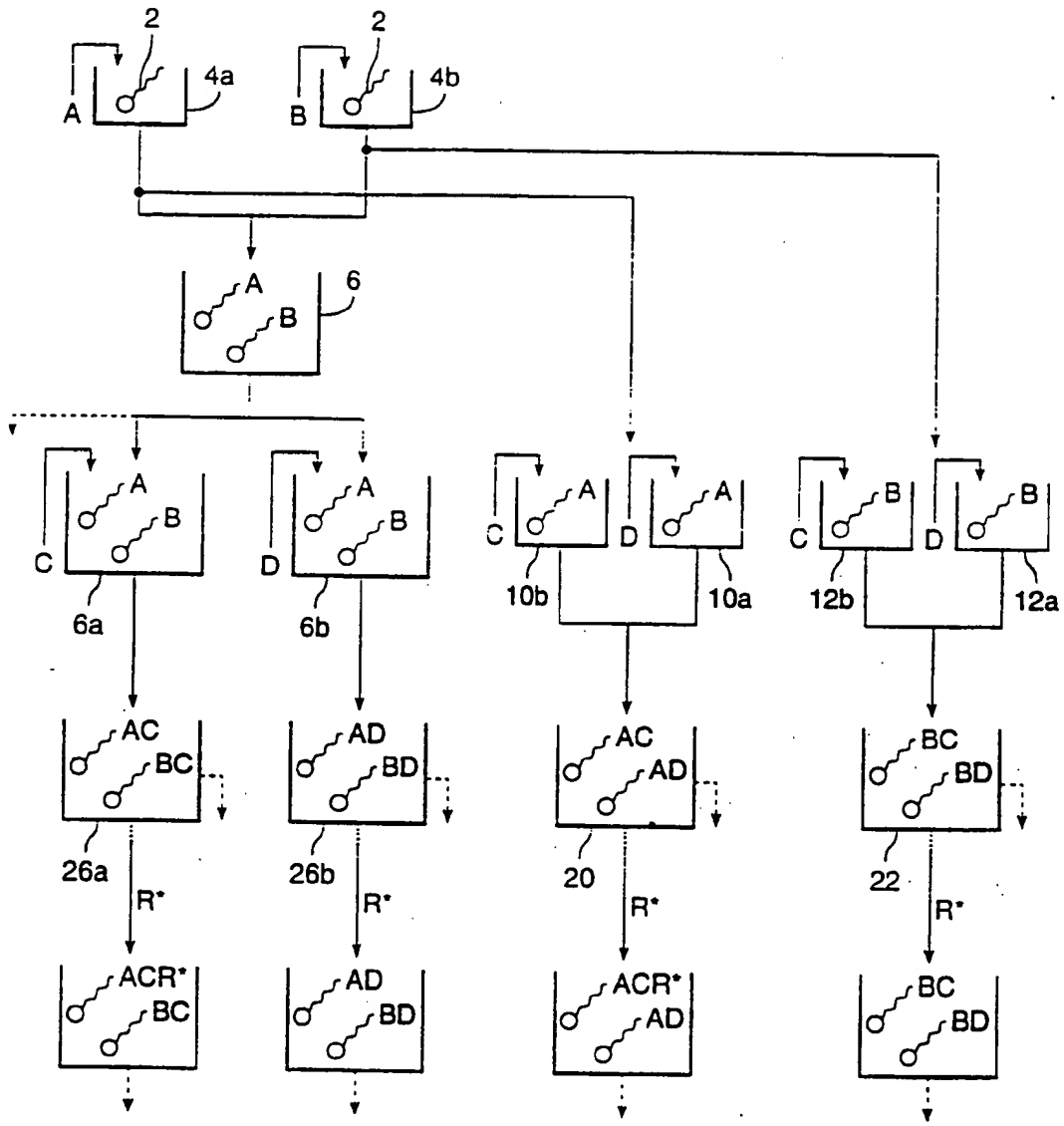
2
2
4a
4b
A
B
A
B
6
C
D
6a
6b
10b
10a
12b
12a
AC
BC
AD
BD
AC
AD
BC
BD
26a
26b
20
22
R*
R*
R*
R*
ACR*
BC
AD
BD
ACR*
AD
BC
BD

FIG. 1a

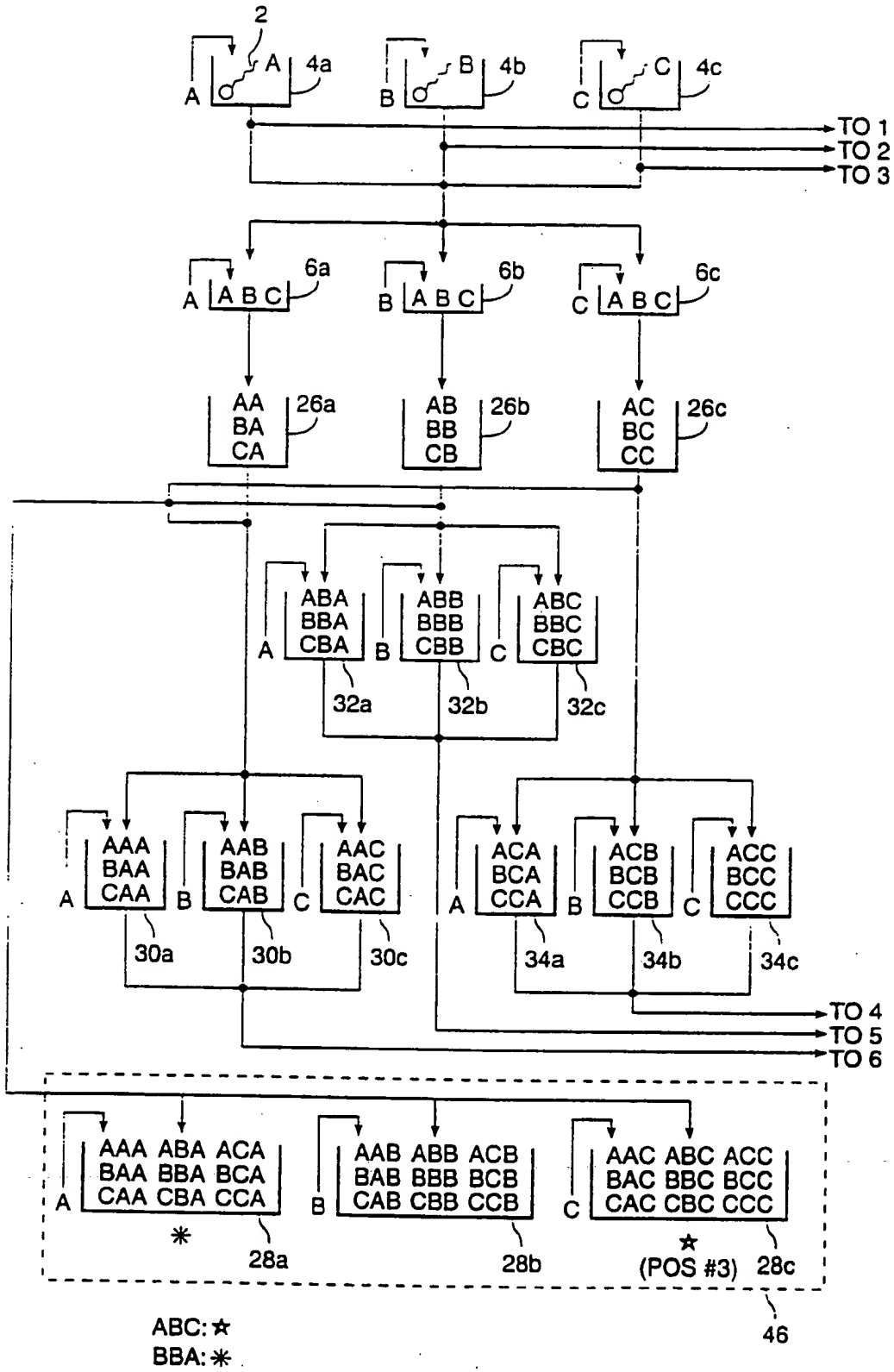
2
A
4a
B
4b
C
4c
TO 1
TO 2
TO 3
6a
6b
6c
A B C
AA BA CA
26a
AB BB CB
26b
AC BC CC
26c
ABA BBA CBA
ABB BBB CBB
ABC BBC CBC
32a
32b
32c
AAA BAA CAA
AAB BAB CAB
AAC BAC CAC
30a
30b
30c
ACA BCA CCA
ACB BCB CCB
ACC BCC CCC
34a
34b
34c
TO 4
TO 5
TO 6
AAA ABA ACA BAA BBA BCA CAA CBA CCA
AAB ABB ACB BAB BBB BCB CAB CBB CCB
AAC ABC ACC BAC BBC BCC CAC CBC CCC
28a
28b
(POS #3)
28c
46
ABC: ☆
BBA: ✳

FIG. 1b.1

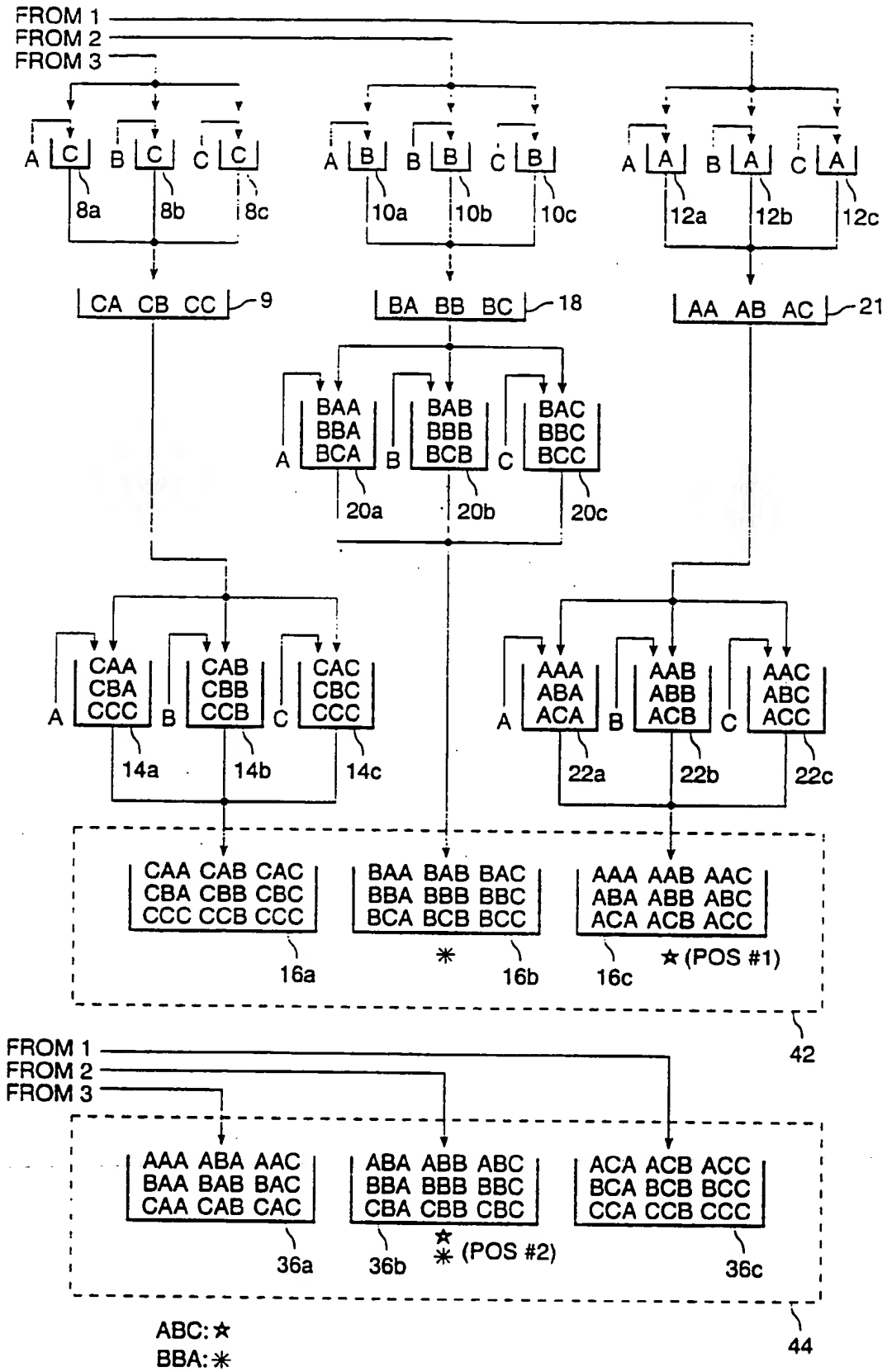
FROM 1
FROM 2
FROM 3
A C B C C C
8a 8b 8c
A B B B C B
10a 10b 10c
A A B A C A
12a 12b 12c
CA CB CC 9
BA BB BC 18
AA AB AC 21
BAA BBA BCA
BAB BBB BCB
BAC BBC BCC
20a 20b 20c
CAA CBA CCC
CAB CBB CCB
CAC CBC CCC
14a 14b 14c
AAA ABA ACA
AAB ABB ACB
AAC ABC ACC
22a 22b 22c
CAA CAB CAC
CBA CBB CBC
CCC CCB CCC
16a
BAA BAB BAC
BBA BBB BBC
BCA BCB BCC
16b
AAA AAB AAC
ABA ABB ABC
ACA ACB ACC
16c
★ (POS #1)
42
FROM 1
FROM 2
FROM 3
AAA ABA AAC
BAA BAB BAC
CAA CAB CAC
36a
ABA ABB ABC
BBA BBB BBC
CBA CBB CBC
36b
☆
✳ (POS #2)
ACA ACB ACC
BCA BCB BCC
CCA CCB CCC
36c
44
ABC: ★
BBA: ✳

FIG. 1b.2

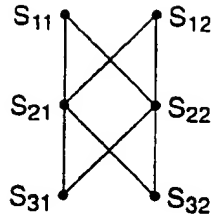

FIG. 2

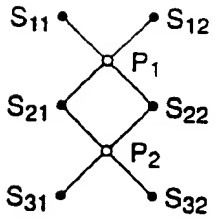

FIG. 3

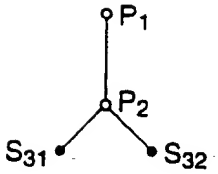

FIG. 4

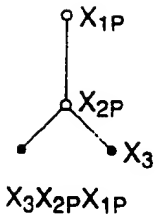

FIG. 5a

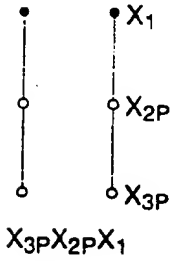

FIG. 5b

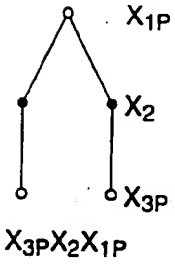

FIG. 5c

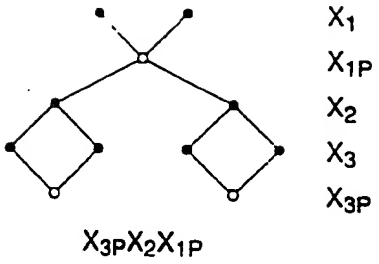

FIG. 6

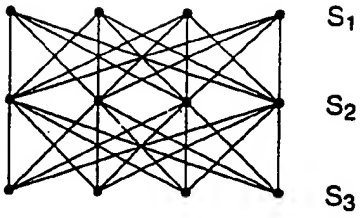

FIG. 7

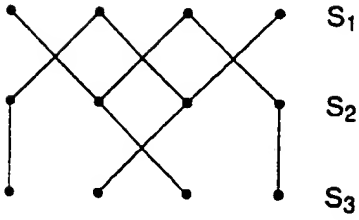

FIG. 8

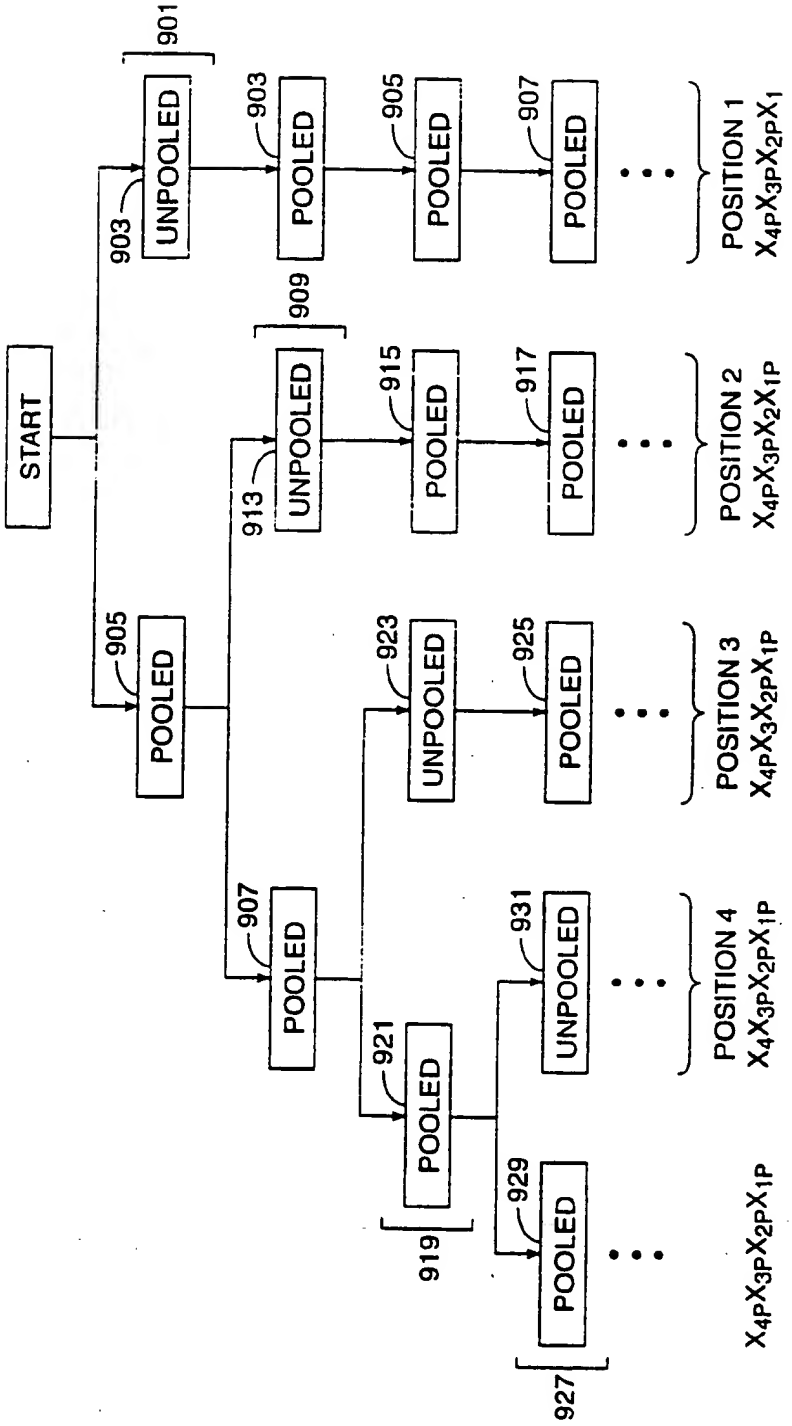
START
UNPOOLED
POOLED
901
903
905
907
909
913
915
917
919
921
923
925
927
929
931
POSITION 1
$X_{4P}X_{3P}X_{2P}X_1$
POSITION 2
$X_{4P}X_{3P}X_2X_{1P}$
POSITION 3
$X_{4P}X_3X_{2P}X_{1P}$
POSITION 4
$X_4X_{3P}X_{2P}X_{1P}$
$X_{4P}X_{3P}X_{2P}X_{1P}$

FIG. 9

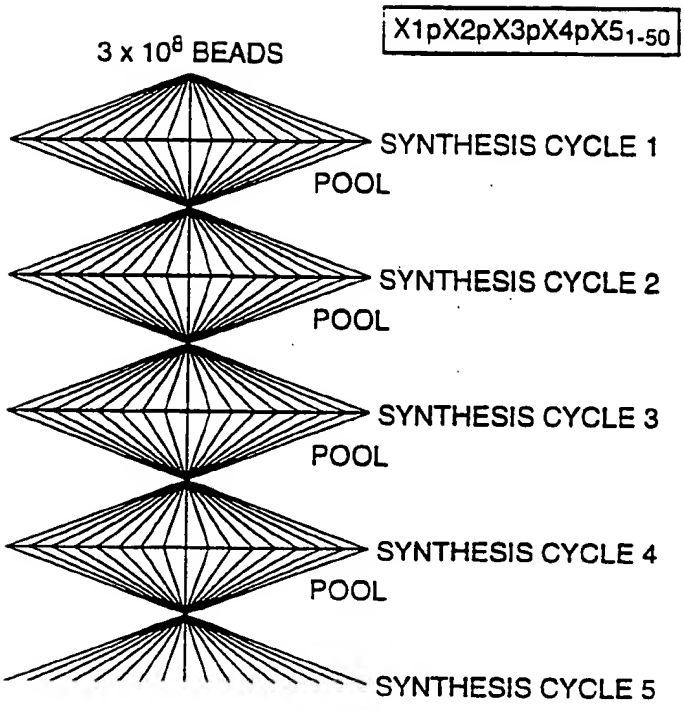

FIG. 10a

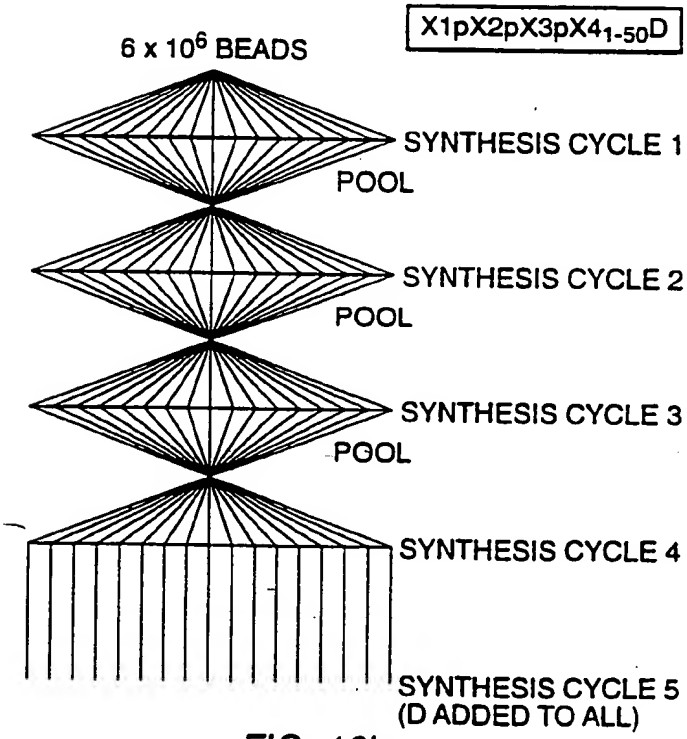

FIG. 10b

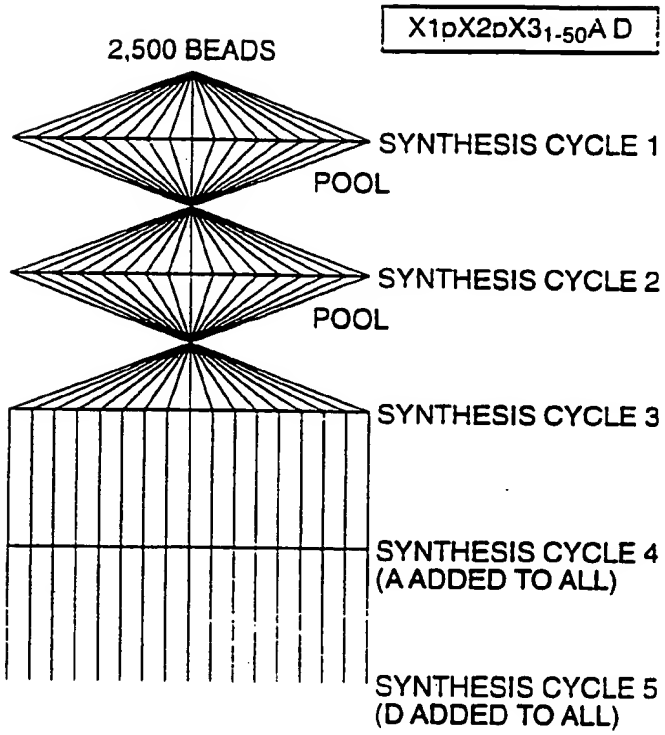

FIG. 10c

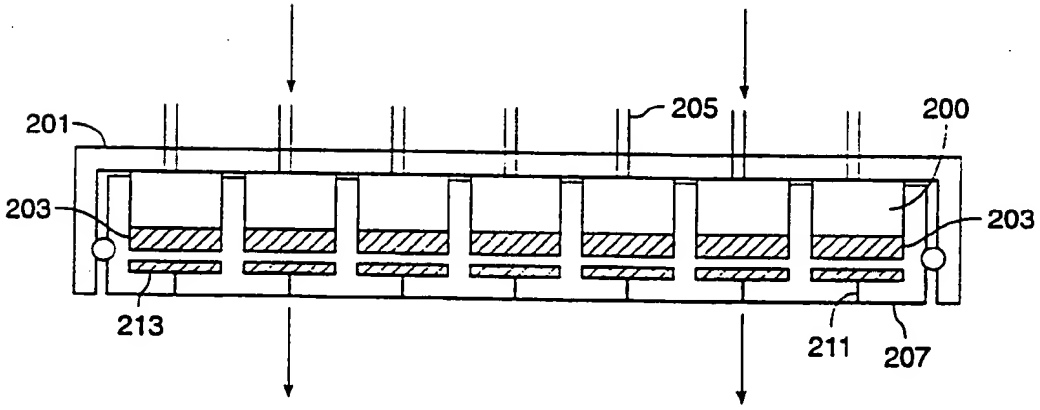

FIG. 11a

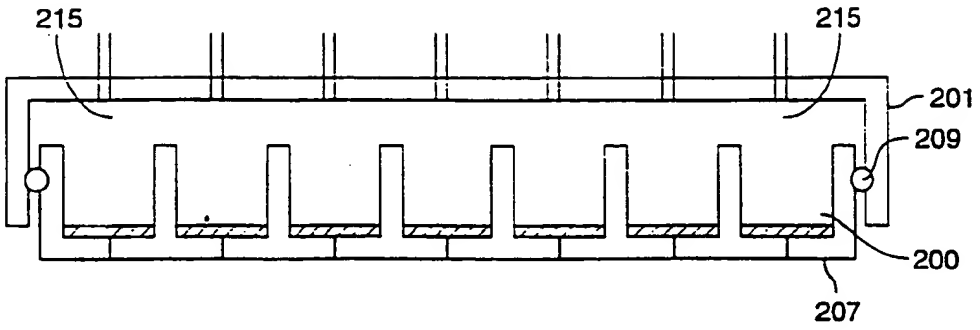

FIG. 11b

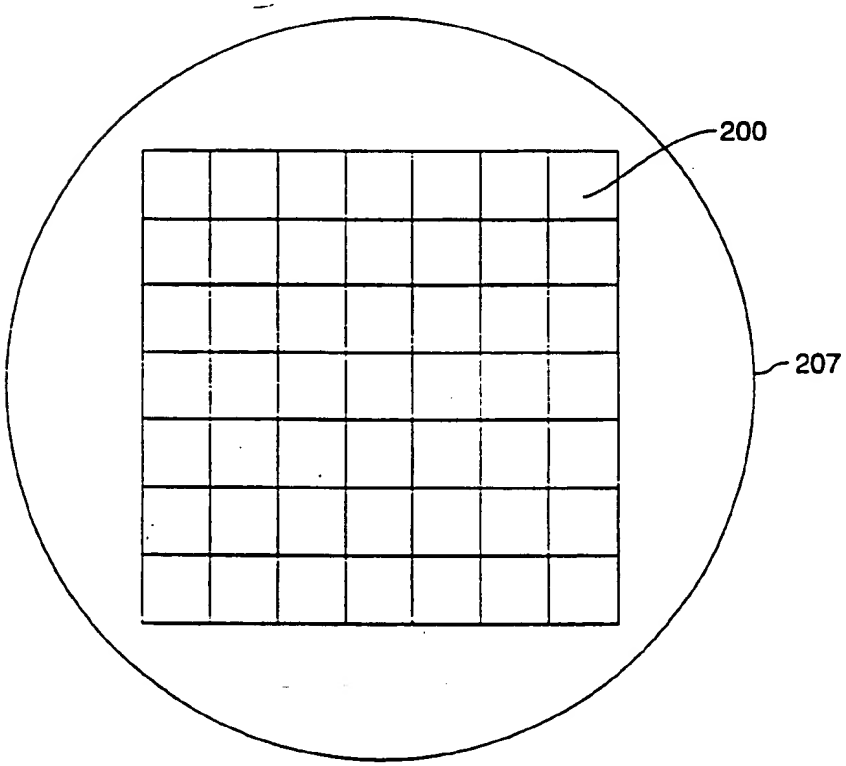

FIG. 11c

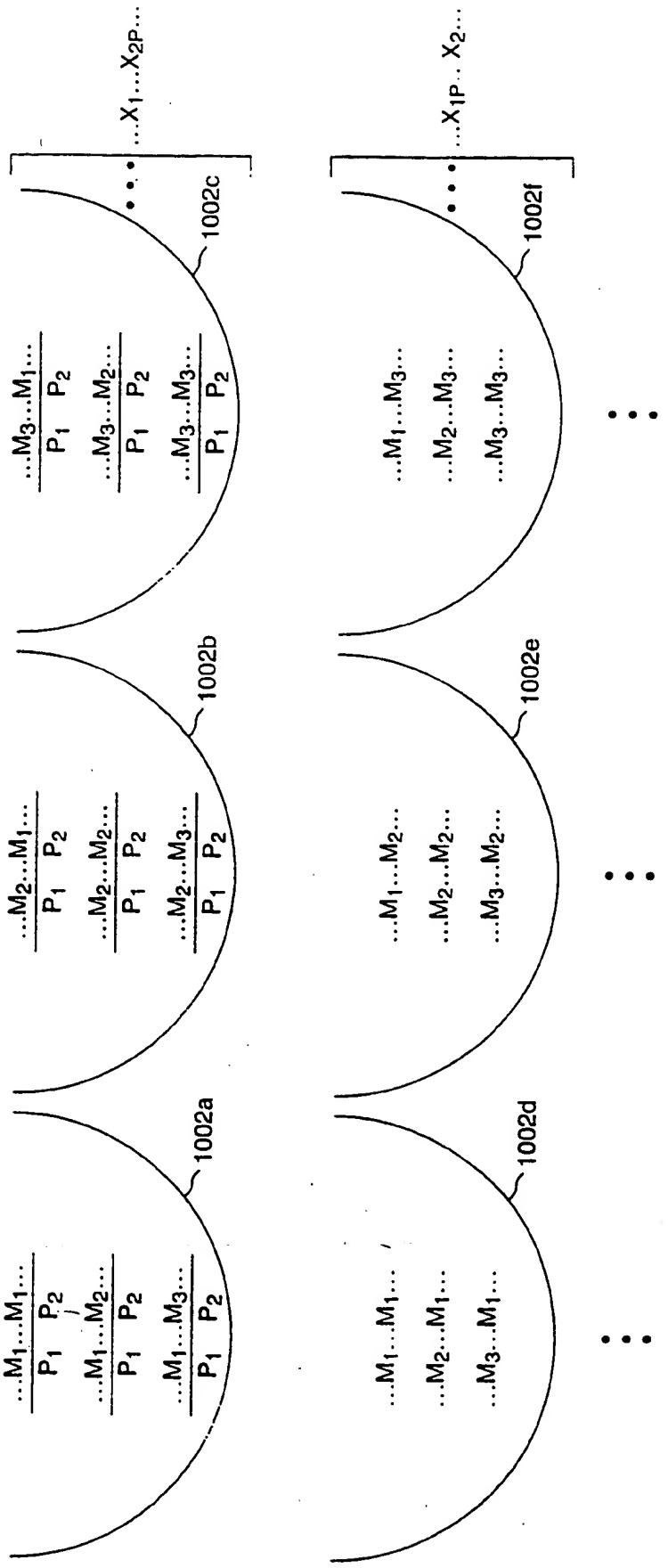
...M1...M1...
P1 P2
...M1...M2...
P1 P2
...M1...M3...
P1 P2
1002a
...M2...M1...
P1 P2
...M2...M2...
P1 P2
...M2...M3...
P1 P2
1002b
...M3...M1...
P1 P2
...M3...M2...
P1 P2
...M3...M3...
P1 P2
1002c
...X1...X2P...
...M1...M2...
...M2...M2...
...M3...M2...
1002e
...M1...M3...
...M2...M3...
...M3...M3...
1002f
...X1P.. X2...
...M1...M1...
...M2...M1...
...M3...M1...
1002d

FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/04145

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :G01N 33/68, 33/543; C07K 17/00; C12Q 1/70, 1/68
US CL :435/5, 6; 436/501, 518; 530/334
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/5, 6; 436/501, 518; 530/334

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN, Dialog, APS
search terms; peptide screening, peptide libraries, combinatorial libraries

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	BioTechniques, Vol. 13, No. 6, issued December 1992, Pinilla et al, "Rapid Identification of High Affinity Peptide Ligands Using Positional Scanning Synthetic Peptide Combinatorial Libraries", pages 901-905, see entire article.	1-17
Y	WO, A, 86/00991, (Geysen)) 13 February 1986, see pages 4-7 and 11-13.	1-17
Y	European Federation for Medicinal Chemistry Xth International Symposium on Medicinal Chemistry ABSTRACTS, published 1988, Furka et al. "More Peptides by Less Labour", P-168, see entire abstract.	1-17

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
10 June 1993

Date of mailing of the international search report
JUL 13 1993

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. NOT APPLICABLE

Authorized officer
LORA M. GREEN

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/04145

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SCIENCE, Vol. 251, issued 15 February 1991, Fodor et al., "Light-Directed, Spatially Addressable Parallel Chemical Synthesis", pages 767-778, see entire article.	5, 14